

Factor analysis of clinical data from asbestos workers: implications for diagnosis and screening

P HARBER, M LEW, D P TASHKIN, M SIMMONS

From the Occupational Medicine Branch and Pulmonary Function Laboratory, Pulmonary Division, Department of Medicine, University of California, Los Angeles, CA 90024, USA

ABSTRACT Clinical data from 624 asbestos exposed workers were analysed by factor analysis. Fifteen clinical variables were found to reflect several underlying factors: "obstruction," "interstitial disorder," "x ray change," "air trapping," and "age/exposure." Scoring individual subjects along these axes can facilitate identifying early cases and avoiding false diagnosis. The analysis suggests that screening should be multimodal and that radiographic abnormality need not imply physiological impairment.

Accurate diagnosis of asbestos related disease depends on the integration of much clinical data. Whereas the legal system may occasionally recognise a single test—radiography for example—as specific and sensitive, knowledgeable clinicians recognise the importance of obtaining several different types of data in each subject. The pulmonary system is complex, having multiple components (large and small airways, interstitial tissue, vasculature, pleural tissue, and respiratory muscle), and clinical information and laboratory tests relate to specific pathological and underlying physiological abnormalities only indirectly. Thus clinical data (individual spirometry test results, symptoms, or radiographs) may be seen as indirect reflections of the underlying process (or factors).

Factor analysis is a technique that may be used to explore the interrelations among observed data to assess underlying factors that may not be directly observable.^{1,2} This technique was used to analyse data from a group of heavily asbestos exposed subjects to assess the nature of underlying factors. The results suggest the presence of several different processes that require different clinical and surveillance approaches.

Methods

Results of clinical evaluations of 624 asbestos exposed workers were analysed. All had been examined at the occupational medicine branch and pulmonary division of the UCLA Medical Center

between 1980 and 1984. The subjects were referred because of possible abnormalities noted in a shipyard screening programme or because of filing a compensation claim. Their mean age was 55.9 ± 10.4 (standard deviation [SD]) years. Only men were included in this analysis; 46.5% were white and 43.3% black. Most shipyard workers and their duration of exposure ranged from one to 51 years (mean 20.8 ± 10.2 [SD] years). Although specific industrial hygiene data were not available, most exposures were moderate to heavy in nature.

Each subject completed a standard questionnaire, was interviewed by a single nurse clinician, and examined by a physician. Spirometry was performed with measurement of forced expiratory volume in one second (FEV₁), forced vital capacity (FVC), the FEV₁/FVC ratio (Ratio), and the forced expiratory flow over the mid-vital capacity range (FEF_{25-75%}). Data were determined from the spirogram with the highest sum of FEV₁ and FVC. Only prebronchodilator values are reported. Spirometry was performed using an automated pneumotachograph spirometer (SRL Series, Gould, Inc, Houston) or a 13.5 l water-seal spirometer (WE Collins, Braintree, MA). Diffusing capacity for carbon monoxide (D_{CO}) was determined by the single breath method³ using the SRL unit or a Collins modular diffusing capacity unit (Braintree). Total lung capacity (TLC), functional residual capacity (FRC), and residual volume (RV) were measured by nitrogen washout⁴ (SRL) or by helium dilution⁵ (Collins Modular) methods. These analyses include the TLC and RV data.

Data for FEV₁ and FVC were adjusted for age and height using the regression equations of Crapo *et al.*⁶ FEF_{25-75%} was adjusted by the equations of

Table 1 Characteristics of population

	All (n = 624)	Current smokers (n = 166)	Non-smokers (n = 146)
Age (years)	55.9 ± 10.4	54.1 ± 9.8	56.0 ± 12.1
Duration of exposure (years)	20.8 ± 10.3	20.0 ± 10.9	20.6 ± 10.1
Race: % white	46.5		
% black	43.8		
FEV ₁ R (l)	- 0.27 ± 0.73	- 0.35 ± 0.78	- 0.01 ± 0.60
FVCR (l)	- 0.49 ± 0.68	- 0.48 ± 0.70	- 0.38 ± 0.65
D _{co} R (ml/mm/min)	- 3.9 ± 5.1	- 5.5 ± 5.0	- 2.4 ± 4.4
FEF _{25-75%} R (l/sec)	- 1.6 ± 1.9	- 1.7 ± 3.2	- 1.2 ± 1.1

Results are shown as mean ± SD. Pulmonary function tests are defined in the text and are shown as residuals (observed-predicted).

Knudsen *et al.*⁷ Cotes's formula was used for adjusting D_{co},⁸ and the equations of Goldman and Becklake were used for adjusting RV and TLC.⁹ As recommended by the recent American Medical Association guidelines,¹⁰ the observed FEV₁ and FVC of blacks were divided by 0.9 before age and height adjustment. The residual (observed minus predicted) was used to express the adjusted values; residual values are denoted by appending "R" to the variable (FEV₁R, FVCR, FEF_{25-75%}R, D_{co}R, RVR, TLCR). An air trapping index (ATI) was represented by the RV/TLC ratio.

Arterial blood, obtained while subjects breathed ambient air, was analysed for oxygen tension (PO₂), using a semi-automated blood gas analyser (Corning, model 168, Medford, MA).

Each subject had posterior/anterior, lateral, and bilateral oblique chest x ray films taken and interpreted by chest radiologists. The radiologists' reports were abstracted and numerically coded. There were separate scores for interstitial (XR-INT) and pleural (XR-PL) abnormalities, assigning higher scores for greater certainty of abnormality and extent of abnormality. In addition, the abstractor assigned a numerical dyspnoea score (1-5) based on the recorded questionnaire and interview data.

Subjects were considered non-smokers if they had

not smoked for over five years and had less than one pack-year of cigarette smoking over their lifetime; they were considered current smokers if they had smoked regularly within six months of examination. The remainder were considered to be ex-smokers.

Data were collected using an HP87 microcomputer (Hewlett-Packard, Corvallis, OR) and transferred to a PDP-11/23 computer (Digital Equipment Corp, Maynard, MA) for subsequent analysis. BMDP statistical programs² were used for analysis. After preliminary univariate analysis of each variable, the P4M program was used for factor analysis.² Initial factor extraction was by principal components analysis. The varimax method was used for factor rotation.^{1,2} Data were standardised before factor identification. To avoid any biases, we decided a priori to accept the BMDP default criteria for factor identification and retention. (A subsequent analysis, forcing retention of exactly four factors, did not yield substantially different results.)

Analysis was performed using all subjects. In addition, separate analyses were performed for the non-smoker and current smoker subgroups.

Results

Table 1 shows the characteristics of the study popula-

Table 2 Factor loadings for all subjects

Factor	1	2	3	4	5
Ratio	0.865	—	—	—	—
FEF _{25-75%} R	0.725	—	—	—	—
FEV ₁ R	0.614	0.570	—	—	—
FVCR	—	0.909	—	—	—
TLCR	- 0.385	0.740	0.324	—	—
D _{co} R	0.345	0.537	—	—	—
RVR	—	—	0.956	—	—
ATI	—	—	0.951	—	—
XR-INT	—	—	—	0.784	—
XR-PL	—	—	—	0.704	—
EXP	—	—	—	—	0.682
AGE	- 0.265	—	0.266	0.264	0.638
CIG	0.350	—	—	—	0.594
DYSP	—	- 0.328	—	—	—
PO ₂ (resting)	0.365	—	—	- 0.434	—
"Title"	Obstruction	Interstitial	Air trapping	x ray	Age exposure

The factor loadings for variables (standardised to mean = 0, SD = 1) are shown. Columns are arranged in order of decreasing variance explained. Loadings < 0.25 have been deleted. Variables are defined in text. Titles are descriptive only.

Table 3 *Factor loading for current smokers*

Factor	1	2	3	4
RVR	0.909	—	—	—
Ratio	-0.824	—	—	—
ATI	0.805	—	—	—
FEVR	-0.617	—	—	—
FVCR	—	0.885	—	—
TLCR	0.588	0.712	—	—
D ₅₀ R	—	0.639	—	—
XR-INT	—	—	0.754	—
XR-PL	—	—	0.728	-0.295
AGE	0.410	—	0.569	0.382
EXP	—	—	0.528	0.256
DYSP	—	—	—	0.836
PO ₂	-0.425	—	—	—
CIG	—	—	—	—
FEF _{25-75%}	0.476	—	—	—
"Title"	Obstruction	Interstitial	x ray/age exposure	Age/pleura dsypnoea

See legend for table 2.

Table 4 *Factor loading for non-smokers*

Factor	1	2	3	4
FVCR	0.922	—	—	—
TLCR	0.854	0.254	—	—
D ₅₀ R	0.611	—	—	—
AGE	—	0.733	—	0.384
ATI	—	0.729	-0.435	—
XR-INT	—	0.567	—	—
FEF _{25-75%}	—	—	0.869	—
Ratio	—	-0.253	0.844	—
FEV ₁	0.573	—	0.657	—
EXP	—	—	—	0.746
XR-PL	—	—	—	0.685
DYSP	-0.385	0.355	—	—
CIG	—	—	—	—
PO ₂	—	-0.453	—	—
"Title"	Interstitial	Air trap/x ray	Obstruction	Pleura/exposure/age

See legend for table 2.

tions. In general, the subjects had had prolonged exposure to asbestos (mean duration = 21 years) and were middle aged (mean age 56). As expected, lung function in current smokers was worse than in non-smokers. Spirometric findings suggestive of restrictive lung disease (FVC < lower bound of 95% confidence interval with FEV₁/FVC ratio ≥ predicted) were present in 4.4% of the population.

Tables 2-4 summarise the factor analyses for all subjects, current smokers, and non-smokers. The factors are listed in order of decreasing variance explained. The relation between each factor and each clinical variable is shown by the correlation between the factor and the variable. For clarity, low values (loading < 0.25) have been deleted from the tables. For convenient reference, we have assigned a descriptive name, or "title" to each factor.

Methods for calculating factor scores (scale values) for all subjects are shown in table 5. Each score—for example, "restrictive abnormality"—is a weighted sum of selected variables, weights given by the factor score coefficients shown in table 5. Using these equa-

tions, each subject's information may be summarised by scores on 4-5 indices, representing different processes.

Discussion

Determining the presence or absence of asbestos related disease and assessing its significance requires integrating multiple data items concerning each subject. This study examined the interrelations among different elements of the clinical data base using a technique requiring minimal a priori clinical assumptions. The analysis has led to several implications about the nature of asbestos effects.

Factor analysis seeks to find underlying factors to explain the observed test values.^{1,2} The descriptive model assumes that the underlying factors—obstruction, interstitial, for example—explain the directly observable variables—for example, FEV₁. The standardised factor loadings represent the closeness of association of the underlying factor with an observable variable. The statistical method seeks to

Table 5 Calculation of a subject's factor scores

Factor score 1 ("obstruction") = $0.0406 (\% \text{ ratio} - 73.98) + 0.1959 (\text{FEFR} + 1.555) + 0.3176 (\text{FEVR} + 0.0268) - 0.2393 (\text{TLCR} + 0.1224) + 0.0240 (\text{D}_{\text{CO}}\text{R} + 3.39) - 0.0064 (\text{age} - 55.95) + 0.231 (\text{CIG}^* - 1.968) + 0.0148 (\text{PO}_2 - 82.93)$

Factor score 2 ("interstitial") = $0.283 (\text{FEVR} + 0.268) + 0.633 (\text{FVRC} + 0.487) + 0.410 (\text{TLCR} + 0.1224) + 0.0425 (\text{D}_{\text{CO}}\text{R} - 3.390) - 0.128 (\text{DYSPT} + 2.657)$

Factor score 3 ("air trap") = $0.0892 (\text{TLCR} + 0.1224) + 0.444 (\text{RVR} - 1.852) + 3.33 (\text{RV/TLC} - 0.3558) + 0.0028 (\text{AGE} - 55.95)$

Factor score 4 ("x ray") = $0.575 (\text{INT-RAD}^\dagger - 1.663) + 0.0749 (\text{PL-RAD}^\S - 15.04) + 0.0062 (\text{AGE} - 55.95) - 0.0318 (\text{PO}_2 - 82.93)$

Factor score 5 ("EXP-AGE") = $0.0506 (\text{EXP} - 20.85) + 0.0427 (\text{AGE} - 55.95) + 0.727 (\text{CIG}^* - 1.968)$.

*CIG: 1, for current smoker; 2, for ex-smoker; 3, for never smoked.

†DYSPT: 1, no dyspnoea; 2, hurrying on level ground or slight hill; 3, walking on level ground at pace of others of own age; 4, walking on level ground at own pace; 5, at rest.

‡INT-RAD: 1, no increased markings; 2, abnormal markings;

3, possible (non-specific); 4, probable; 5, moderate; 6, severe.

§PL-RAD: 10, no thickening; 20, local thickening; 30, diffuse thickening.

The factor scores for an individual are calculated from the observed variables as shown. The factor score coefficients allow use of "raw" data rather than results converted to "standard scores" (mean = 0, SD = 1). Coefficients are derived from all subjects.

identify factors that are not closely related. The technique has been extensively used in the social sciences.

We analysed the entire study population and two subgroups defined by smoking status (lifelong non-smokers, current smokers). The ex-smokers group was not analysed separately because of a lack of homogeneity—for example, differences in interval since stopping smoking and reasons for smoking cessation. As shown in tables 2–4, fairly consistent results were found. The factors noted relate to the clinical tests in a manner suggesting their clinical meaning. One factor, which we have termed interstitial, relates closely to the physiological tests of lung size, such as TLC, and gas exchanging ability— D_{CO} , for example. Others might label this factor "restrictive" but we avoided this term since "restrictive physiological abnormality" describes a lung size without regard to gas exchanging ability. A second factor, termed obstruction, appears related to airway function. The "air trapping factor," identified in the all subjects and non-smoker groups reflects incomplete lung emptying, possibly due to airway closure during the spirometric forced expiratory manoeuvre.¹¹ Aging and cigarette smoking may lead to this effect. It is, therefore, not surprising that ATI is associated with age and does not appear as an independent factor in the current smokers group, wherein smoking causes damage leading to both air-flow obstruction and air trapping. In addition to reflecting air trapping, a raised RV/TLC ratio (ATI) might theoretically reflect interstitial disease if ability to inspire maximally is more affected than the residual volume. It appears

here, however, that the ATI does reflect air trapping since it does not appear to be closely associated with other interstitial disease variables. In the all subjects and non-smoker groups both radiographic variables (reflecting pleural and interstitial changes) were the predominant variables in a separate factor that we have labelled "x ray." Age and its covariate, exposure, also appeared in these factors. In both current smokers and non-smokers but surprisingly not in the all subjects group age, exposure, and pleural changes appeared related to a common factor.

Use of the factor score coefficients to calculate scores for each factor for each person can help simplify clinical assessments by summarising 15 elements in 4–5 scores. Furthermore, this method provides a "rational" means for combining data. Thus an isolated minor abnormality of a single test may not in itself imply "disease" if the subject's factor scores (combined from multiple tests) are all normal. Conversely, use of factor scores may permit identification of disease even if each individual test result is technically within the "normal range." Early studies by Murphy *et al* showed that reliance on any single pulmonary function variable is inadequate for diagnosing asbestosis.¹²

These findings may be interpreted to suggest that several distinct underlying processes are measured by clinical assessments. The factors identified are "biologically plausible" since a pattern consistent with other respiratory physiological processes was shown.

The statistical method attempted to delineate uncorrelated, independent underlying factors. (Principal component analysis was used for initial factor extraction.¹) The findings suggesting several distinct factors have several possible implications. Firstly, there is a potential means to separate asbestos from non-asbestos effects. Effects of asbestos on lung volumes and on interstitial tissue are well recognised. Although many reports recently summarised by Becklake and Ernst suggest that asbestos may also produce airway effects,¹³ these are pathologically and physiologically located in small airways^{14 15} and may be quantitatively small by comparison with the effect of cigarette smoking. Thus, particularly in smokers, the abnormalities of scores on the obstruction axis are much less likely to be attributable to asbestos effects than are abnormalities on other axes, such as interstitial.

Secondly, radiographic findings, particularly in the analysis of all subjects, appeared related to separate factors from those most associated with the physiological test variables. Thus these data support the view that radiographic changes may not be closely related to physiological impairment, and the presence of mild radiographic abnormalities need not imply physiological impairment. Weill *et al* have also noted the dissociation between physiological and radiographic manifestations of asbestos exposure.¹⁶

Thirdly, screening programs should recognise the multifaceted nature of asbestos effects. If there are actually several poorly related lung processes that might be affected by asbestos screening efforts must be directed at each. Therefore, chest radiography, spirometry, and perhaps diffusing capacity testing may additively increase sensitivity. Other investigators, using other analytical methods, have also suggested that the initial manifestation of asbestos effect may differ between individuals.¹²

It is not surprising that age and exposure appeared in the same factor since older workers here had more years to accumulate exposure. Pleural plaque radiographic score was also associated with this underlying factor, suggesting that plaque formation may be associated with exposure and aging, that plaques become more apparent with age, or that they may simply be a manifestation of dose and latency. The epidemiological studies of Becklake *et al* have also shown a close association between age and plaques.¹⁷

The conclusions suggested by the present study must be considered tentative. Multivariate techniques such as factor analysis are subject to probabilistic variation and may be affected by choice of specific analytical technique. We attempted, however, to avoid introducing any bias by intentionally choosing the "default" parameters of BMDP.²

The population we studied may not be typical of workers currently exposed to asbestos. Their exposures were heavy and extended over many years and most of our subjects were in an older age group. Although many of the radiographs were not read by the ILO system,¹⁸ the radiologists were unaware of the clinical information about each subject, and therefore any systematic radiographic bias is unlikely.

Other workers have used "scales" to combine and summarise clinical data from asbestos exposed workers. The criteria of Murphy *et al* for the diagnosis of asbestosis required several abnormalities for the diagnosis.¹² These criteria, which are still commonly used, characterised each variable dichotomously as "abnormal" or "normal"; therefore, a subject with "low normal" values for all tests would not be detected as diseased. Fournier-Massey and Becklake used an arbitrary scale to place each subject along a continuum from restrictive to normal to obstructive abnormality.¹⁹ Use of multiple scales, each related to a separate asbestos effect, may however be a more appropriate means of summarising clinical data. Regan *et al* used a multivariate technique to examine pulmonary function data from a group of 201 asbestos exposed workers.²⁰ Their principal components analysis yielded 16 components, the first of which separated "lung disease from health," and the second of which was inter-

preted as separating asbestosis from obstructive disease.

We thank Karen SooHoo for help in collecting data and Nancy Marshello for preparing the manuscript.

References

- 1 Afifi AA, Clark V. *Computer-aided multivariate analysis*. Belmont, CA: Lifetime Learning Publications, 1984.
- 2 Frane J, Jennrich R, Sampson P. Factor analysis. In: Dixon WJ, ed. *BMDP statistical software*. Berkeley: University of California Press, 1983.
- 3 Ogilvie CM, Forster RE, Blakemore WS, Morton JW. A standardized breathholding technique for the clinical measurement of the diffusing capacity of the lung for carbon monoxide. *J Clin Invest* 1957;36:1-17.
- 4 Darling RC, Cournand A, Richards DW. Studies on the intrapulmonary mixture of gases. III. An open circuit method for measuring residual air. *J Clin Invest* 1940;19:609-18.
- 5 Meneely GR, Ball CY, Kory RC, *et al*. A simplified closed circuit helium dilution method for the determination of residual volume of the lungs. *Am J Med* 1960;28:824-31.
- 6 Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meets ATS recommendations. *Am Rev Respir Dis* 1981;123:450-60.
- 7 Knudson RJ, Slatin RC, Lebowitz MD, Burrows B. The maximal expiratory flow-volume curve. Normal standards, variability and effects of age. *Am Rev Respir Dis* 1976;113:587-600.
- 8 Cotes JE. *Lung function*. Oxford: Blackwell, 1975:241-59.
- 9 Goldman HI, Becklake MR. Respiratory function tests. Normal values at median altitudes and the prediction of normal results. *American Review of Tuberculosis* 1959;79:457-67.
- 10 American Medical Association. *Guides to the evaluation of permanent impairment*. Chicago: AMA, 1984.
- 11 Cotes JE. *Lung Function*. 4th ed. London: Blackwell, 1979:411.
- 12 Murphy RHL, Gaensler EA, Ferris BG, Fitzgerald M, Solliday N, Morrissey W. Diagnosis of "asbestosis", observations from a longitudinal survey of shipyard pipe coverers. *Am J Med* 1978;65:488-98.
- 13 Becklake M, Ernst P. Asbestos exposure and airway responses. In: Gee JBL, ed. *Occupational lung disease*. New York: Churchill-Livingstone, 1984.
- 14 Wright JL, Churg A. Morphology of small-airway lesions in patients with asbestos exposure. *Hum Pathol* 1984;15:68-74.
- 15 Bégin R, Cantin A, Berthiaume Y, Boileau R, Pélouguin S, Massé S. Airway function in lifetime-nonsmoking older asbestos workers. *Am J Med* 1983;75:631-8.
- 16 Weill H, Ziskine MM, Waggenspack C, Rossiter CE. Lung function consequences of dust exposure in asbestos cement manufacturing plants. *Arch Environ Health* 1975;30:88-97.
- 17 Becklake MR, Thomas D, Liddell F, McDonald JC. Follow-up respiratory measurements in Quebec chrysotile asbestos miners and millers. *Scand J Work Environ Health* 1982;8(suppl): 105-10.
- 18 International Labour Office. *Guidelines for the use of ILO international classification of radiographs of pneumoconioses*. Geneva: ILO, 1980. (ILO occupational safety and health series No 22 (rev 80).)
- 19 Fournier-Massey GF, Becklake MR. Pulmonary function profiles in Quebec asbestos workers. *Bull Physiopathol Respir* 1978;11:429-45.
- 20 Regan GM, Tagg B, Walford J, *et al*. The relative importance of clinical radiological and pulmonary function variables in evaluating asbestosis and chronic obstructive airway disease in asbestos workers. *Clin Sci* 1971;41:569-81.